DOCUMENT RESUME

ED 234 095 TM 830 608

AUTHOR Siegel, Arthur I.
TITLE Trainability Testing for Navy Selection and Classification.
INSTITUTION Applied Psychological Services, Inc., Wayne, Pa.
SPONS AGENCY Navy Personnel Research and Development Center, San Diego, Calif.
REPORT NO NPRDC-TR-83-25
PUB DATE Jul 83
CONTRACT N00123-77-C-1369
NOTE 49p.
PUB TYPE Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Classification; Evaluation Methods; Military Personnel; *Performance Factors; *Personnel Evaluation; *Personnel Selection; *Predictive Validity; *Test Construction
IDENTIFIERS Armed Services Vocational Aptitude Battery; *Miniature Training and Evaluation Battery; *Navy

ABSTRACT

This is the third report on the use of a technical classification assessment center (TCAC) for evaluating general detail (GENDET) personnel in the Navy. The first two reports described the development of the TCAC and the validation of the TCAC against supervisors' ratings of job performance for a small exploratory sample. This report presents the further development and modification of the tests and the validation of their usefulness for assigning GENDETS as seamen, firemen, and airmen. A ten-exercise miniature training and evaluation battery (MTEB) was administered to 1,034 Navy recruits who were ineligible for assignment to Navy schools. Followup data were collected after 9 and 18 months. The MTEB exercises were judged to be at an appropriate difficulty level, to possess acceptable discriminating power, and to be adequately unidimensional. The predictive validity of the exercises was found to be moderate. An increase in predictive power was evidenced if the MTEB exercises were employed in conjunction with the Armed Services Vocational Aptitude Battery. The results are interpreted to support the value of the MTEB exercises developed for this study and the MTE concept itself. The appendices include a technical performance criterion form for seamen and a form for collecting criterion data about recruit adjustment. (Author/PN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED234095

NPRDC TR 83-25 JULY 1983

TRAINABILITY TESTING FOR NAVY SELECTION AND CLASSIFICATION

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Navy Personnel Research and Development Center
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER
San Diego, California 92152

NPRDC TR 83-25

July 1983

TRAINABILITY TESTING FOR NAVY SELECTION AND CLASSIFICATION

Arthur I. Siegel

Applied Psychological Services, Inc.
Wayne, Pennsylvania 19087

Reviewed by
Martin F. Wiskoff

Released by
J. W. Renard
Commanding Officer

Navy Personnel Research and Development Center
San Diego, California 92152

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NPRDC TR 83-25	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) TRAINABILITY TESTING FOR NAVY SELECTION AND CLASSIFICATION		5. TYPE OF REPORT & PERIOD COVERED Final Mar 1979 - Mar 1981 6. PERFORMING ORG. REPORT NUMBER 12-08-82
7. AUTHOR(s) Arthur I. Siegel		8. CONTRACT OR GRANT NUMBER(s) N00123-77-C-1369
9. PERFORMING ORGANIZATION NAME AND ADDRESS Applied Psychological Services, Inc. Science Center Wayne, PA 19087		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS ZF55.521.030.01.01
11. CONTROLLING OFFICE NAME AND ADDRESS Navy Personnel Research and Development Center San Diego, California 92152		12. REPORT DATE July 1983 13. NUMBER OF PAGES 48
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Test development	Equal employment opportunity	Evaluation
Selection	Culture fair testing	Assessment
Classification	Armed Services Vocational	Measurement
Test concepts	Aptitude Battery (ASVAB)	Appraisal

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The miniature training and evaluation (MTE) approach to personnel selection involves training a person on a sample of the tasks of a job and then testing him or her on task performance. People who show ability to learn to perform a sample of a job's tasks are assumed to be able to learn the whole job, given appropriate on-the-job training. Such tests are held to meet EEOC requirements better than the usual paper-and-pencil selection tests.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

A ten-exercise MTE battery (MTEB) was administered to 1034 Navy recruits who were ineligible for assignment to Navy schools. Followup data were collected after 9 and 18 months. The MTEB exercises were judged to be at an appropriate difficulty level, to possess acceptable discriminating power, and to be adequately unidimensional. The predictive validity of the exercises was found to be moderate. An increase in predictive power was evidenced if the MTEB exercises were employed in conjunction with the ASVAB. The results are interpreted to support the value of the MTE exercises developed for this study and the MTE concept itself.

FOREWORD

This research was conducted in support of exploratory development work unit ZF55.521.030.01.01 (Prediction of Performance). This is the third report on the use of a technical classification assessment center (TCAC) for evaluating general detail (GENDET) personnel. The first two reports (NPRDC TR 77-3 and TN 82-23) described the development of the TCAC and the validation of the TCAC against supervisors' ratings of job performance for a small exploratory sample. This report presents the further development and modification of the tests and the validation of their usefulness for assigning GENDETS as seamen, firemen, and airmen.

The results of this study are primarily intended for officials in the Navy Military Personnel Command and classification personnel in recruit training commands. The contracting officer's technical representative was Dr. Charles H. Cory.

J. W. RENARD
Commanding Officer

JAMES W. TWEEDDALE
Technical Director

SUMMARY

Problem

Current methods for classifying Navy enlisted personnel are based, to a large extent, on Armed Services Vocational Aptitude Battery (ASVAB) scores. There is evidence, however, that ASVAB scores may be less accurate for classifying enlisted personnel who are likely to receive general detail (GENDET) assignments than for classifying personnel who are eligible for Navy schools. Consequently, new testing methods are needed to determine the assignment of personnel to GENDET jobs. Such testing methods, if available, could lead to improved use of the Navy's limited manpower.

Objective

The purpose of this work was to determine the value of a new testing concept--miniature training and evaluation (MTE) testing--for predicting the fleet performance of GENDET personnel.

Approach

The MTE concept involves training personnel to perform sample tasks and then testing their performance. The concept holds that, regardless of ASVAB scores, recruits who can learn to perform a sample of the tasks of a Navy rating will be able to learn and perform all of the tasks of the rating--given appropriate on-the-job training.

A test battery encompassing the MTE concept was developed and administered to a sample consisting of 1034 recruits at the Naval Training Center, San Diego. The battery contained 10 training-evaluation situations that derived from analysis of typical entry level tasks for seamen, firemen, and airmen. All members of the sample had ASVAB scores below the level required for assignment to any technical training school.

Performance rating data were collected after the recruits had been on their first assignment for about 9 months and, again, about 9 months later. Several statistical analyses were conducted to determine the internal psychometric characteristics of the MTE exercises and the relationship between scores on the miniature training and evaluation battery (MTEB) and the fleet performance ratings.

Findings

1. Correlations between fleet performance ratings and scores on individual ASVAB subtests, and between performance ratings and scores on individual MTE items, were low. This was attributed to problems in the performance rating data. When the correlations were corrected for criterion unreliability, better predictions were obtained.

2. Moderate to high correlations were found between performance ratings and the overall MTEB and ASVAB scores. Accordingly, it seems that the MTEB and ASVAB both predict fleet performance ratings with some degree of success.

3. At the first follow-up (9 months), employing the MTEB scores in conjunction with ASVAB scores more than doubled the predictive power over that obtained with the ASVAB alone. At the second follow-up (18 months), the advantage of the MTE tests was not quite so strong; only about one-third was added to the predictive power already available from the ASVAB.

4. MTEB was consistently superior to the ASVAB for predicting the fleet performance of the sample involved.

Conclusions

This research supports the use of the MTE approach as a technique for augmenting the power of the current classification process for personnel assigned as GENDETs. The MTE tests had generally higher validities than the ASVAB tests for predicting the fleet performance criteria and, when employed in conjunction with the ASVAB, resulted in an increase in the ability to predict performance.

Recommendation

Although MTE variables provided increases in validity coefficients over ASVAB variables, these increases were not sufficient to warrant evaluation of the cost effectiveness of using them for classification and assignment of GENDETs. Therefore, additional evaluation of the validity of MTE variables for this sample will be made using job history variables as criteria.

Subsequent to the validation of MTE variables against job history criteria, the usefulness and cost effectiveness of using MTE variables for classification and assignment of GENDETs will be evaluated. As a part of this evaluation, recommendations will be made concerning future research with and use of MTE variables for GENDET personnel.

CONTENTS

	Page
INTRODUCTION	1
Problem	1
Objective	1
Background	1
The Trainability Concept	1
Miniature Training and Evaluation (MTE) Tests	1
Related Work	2
APPROACH	3
Phase I--Preparatory Steps	3
Review of Literature and Development of Exercises	3
Administrator's Manual and Facilities	6
Administrator Training, Pretest, and Revision	7
Phase 2--Administration	7
Personnel	7
Initial Briefing of the Recruits	7
Training	8
Exercise Completion Time	9
Phase 3--Validation	9
RESULTS	12
Mean and Standard Deviations of Each Exercise	12
Score Distributions	12
Intercorrelations Among Exercises	13
Factor Analysis of MTEB	13
Discussion of Internal Characteristics	15
Predictive Validity Analyses	15
Response to Follow-ups	16
Product-moment Correlations	16
Discussion	20
Correction for Criterion Unreliability	21
Multiple Prediction	21
Comparison of the MTEB with the ASVAB Test	21
Comparison of Multiple Relationships	27
What the MTE Approach Adds	27
DISCUSSION	29
Feasibility of the MTE Approach	29
Future Trends	30
CONCLUSIONS	30
RECOMMENDATION	30
REFERENCES	33

Page

APPENDIX A--TECHNICAL PERFORMANCE CRITERION FORM FOR SEAMEN B-0

APPENDIX B--FORM FOR COLLECTING CRITERION DATA ABOUT RECRUIT ADJUSTMENT C-0

DISTRIBUTION LIST

LIST OF TABLES

1. Exercise Administration Time and Type of Training Included 8

2. Number of Responses for Each Follow-up. 10

3. Maximum Number of Cases Available for Analysis. 11

4. Exercise Means, Standard Deviations and Maximum Possible Range 12

5. Intercorrelations Among Total Test Scores on Miniature Training and Evaluation Exercises 13

6. Results of Factor Analysis of 48 Variables 14

7. Intercorrelations Among Criteria 17

8. Correlations Between MTE Predictors and Criteria 18

9. Correlations Between Criterion Ratings Received in the Two Follow-ups. 20

10. MTE Exercise Validity Coefficients Corrected for Criterion Unreliability. 22

11. Multiple Correlations Between Criteria and MTE Predictors 23

12. Correlation Coefficients Between ASVAB Predictors and Criteria 24

13. Median (M) and Best (B) Uncorrected Correlation for MTEB and ASVAB for SRAT and SPROF Criteria 26

14. ASVAB Validity Coefficients Corrected for Criterion Unreliability 28

15. Multiple Correlations Between ASVAB and Follow-up Criteria 29

16. Multiple Correlation (R) and Percentage of Contribution to Total Predictable Variance (PC) for SRAT Criterion. 29

INTRODUCTION

Problem

Current methods for classifying enlisted Navy personnel are based, to a large extent, on their Armed Services Vocational Aptitude Battery (ASVAB) scores. There is evidence, however, that the ASVAB tests may be less accurate for classifying enlisted personnel who are likely to receive general detail (GENDET) assignments than they are for classifying personnel who are eligible for Navy schools. Consequently, new testing methods are needed to determine the assignment of personnel to GENDET jobs. Such testing methods, if available, could lead to improved use of the Navy's limited manpower.

Objective

The purpose of this work was to determine the value of a new testing concept--miniature training and evaluation (MTE) testing--for predicting the fleet performance of GENDET personnel.

Background

The Navy expects that GENDET recruits will receive substantial on-the-job training before they are assigned to perform difficult tasks in their rating. They will not be expected to perform on their own until after a training or "apprentice" period. Accordingly, the Navy needs to know how trainable individual GENDETs are for specific ratings and if, given adequate training, they can learn to perform competently on the job.

The Trainability Concept

In selecting personnel for job assignments, the predictor (test) attempts to predict the candidate's ability to (1) learn the job (intermediate criterion) and (2) perform on-the-job (ultimate criterion). Accordingly, the Navy must place emphasis on selecting persons who can learn to perform their assigned jobs. Traditional performance or work-sample tests are not practical because recruits cannot be expected to possess the knowledge and skills they will need in some future Navy assignment.

Miniature Training and Evaluation (MTE) Tests

One solution to this problim lies in the MTE approach. In this approach, as described by Siegel and Bergman (1975), the person under consideration is trained to perform a sample of the tasks of the job for which he or she is an applicant and, immediately following the training, his or her ability to perform these tasks is measured. During the training period, full attention is given to individual differences, "hands-on" training, minimization of literacy requirements, individual instruction, etc. The variable evaluated is the performance of the task for which the individual was trained.

For example, one test situation in the MTE battery (MTEB) requires personnel to be taught to start up and shut down a motor and pump apparatus. This involves a 33 step procedure, including a number of safety precautions. The training provides an opportunity for practice. During the test situation, the recruit is asked to start up and then to shut down the apparatus. Scoring is based on adherence to correct procedures and observance of safety precautions.

The approach has a number of advantages. First, it minimizes emphasis on the ability to read and write. This feature answers the objections of these who maintain that written

tests are inherently biased against persons who have less formal education. For a motivated person, formal education, which yields facility with written English, may have little to do with job success in many types of occupations.

Second, because the tasks included within an MTEB are drawn from the actual job to be learned, the content validity of such measures is defensible. This thinking was earlier proposed by O'Leary (1973), who wrote:

> From the validity standpoint, the way out of the dilemma lies in the content validity of the test, or the similarity between test and job. The more nearly the test duplicates the specific tasks to be performed on the job, the greater the chances of developing devices that are fair. Where possible, then, job simulation tests (e.g., job sampling) should be part of the selection procedure. Obviously, this approach has two disadvantages: (a) it is initially more costly than less ideal approaches since a relatively large amount of money must be spent to develop and implement it; and (b) many jobs are so structured that the job sample method is difficult to apply (e.g., testing a 22-year old college graduate for a sales position in a stock brokerage firm). (p. 148)

However, the trainability approach goes beyond performance testing. The training aspect of MTE provides individuals with an opportunity to do well in a way that a performance test without a learning phase does not.

Related Work

The trainability approach to evaluation was first investigated in the Navy context in 1972 (Siegel & Bergman, 1972). A battery of seven training and evaluation situations was constructed for the machinist's mate (MM) rating and administered to 99 recruits who were subsequently assigned as MM strikers in the fleet. A set of performance test criterion instruments was administered after the recruits were on fleet assignment for 9 and 18 months (Siegel & Bergman, 1973; Siegel & Leahy, 1974). The performance criterion in both follow-ups involved a battery of seven individually administered job performance tests. The basis for the criterion instrument set was supervisory opinion relative to an adequately diversified set of tasks that would represent the range of tasks performed by journeyman MMs on the job.

Statistically significant correlations were produced for three of the five criteria at the first follow-up but none were statistically significant at the second follow-up. Comparison of the MTE approach with the Navy's paper-and-pencil tests in use at the time indicated that the MTE approach was superior at the first follow-up, but not at the second.

The MTE concept was later used in an assessment center context (Siegel & Wiesen, 1977). The use of an assessment center to select personnel for technical jobs represents an elaboration of an approach that had been suggested previously by Bray and Moses (1972). Siegel and Wiesen developed a battery of nine MTE exercises and administered them to 140 recruits whose scores on paper-and-pencil tests made them ineligible for assignment to Navy "A" schools. The recruits were evaluated by a team of assessors in terms of how well they would succeed in a selected rating. This allowed the actual MTE exercise scores to be compared with the assessors' combined overall opinion. To this end,

a stepwise multiple regression analysis procedure was used that allowed evaluation of MTE scores relative to the assessors' overall decision. In other contexts, the approach has been termed "policy capturing."

Six stepwise multiple correlation coefficients were calculated between the MTE situations and the assessors' composite estimate for a general rating subgroup, for an MM subgroup, and for the total sample. The resultant multiple correlation coefficients were .68, .81, and .65 respectively, compared to .12, .41, and .23 for the Navy paper-and-pencil tests employed at the time.

Cory (1982) studied the Siegel and Wiesen group after its members were assigned to duty in the fleet using the following criteria: (1) performance ratings by supervisors after the recruits were on the job approximately 1 year and 1-1/2 years, (2) striker/nonstriker status, and (3) retention/attrition. Cory found that MTE scores "had substantial predictive relationships for job performance criteria" and that the MTE scores "added .12 to .31 to the validity coefficients of the maximally predictive batteries of operational variables."

Prior work on the trainability approach done in the United Kingdom, which was reviewed by Robertson and Downs (1979), was primarily concerned with skilled and semiskilled trades in the civil sphere. The criterion in almost all cases was some type of training success. However, remarkably high validity coefficients were generally reported. In the two cases in which an on-the-job rating measure was employed, validity coefficients of .69 and .45 were reported.

The early Navy work with MTE was oriented largely towards developing the methodology. Later work considered the predictive validity of the method but employed only a limited number of Navy ratings and relatively small samples. The present study sought to extend the prior Navy work by investigating the value of the MTE concept for larger samples and a greater number of ratings. As such, the work flowed from the recommendation of Cory (1982), who advised that predictive validation of the method should be carried out on a larger, more definitive sample.

APPROACH

The steps included in the research are shown in Figure 1 and described below.

Phase I--Preparatory Steps

Review of Literature and Development of Exercises

Siegel and Wiesen (1977) had previously developed a battery of training and evaluation situations. Since each situation was, by and large, administered individually, requiring considerable administrative time, it was determined that the battery should be revised to allow a 1/2-day testing period for group administration. The four objectives in the redesign of the battery components were to (1) establish a high degree of correspondence with actual tasks of Navy personnel, (2) minimize paper-and-pencil aspects, (3) establish hands-on training where possible, and (4) create an environment where individualized instruction and feedback are possible.

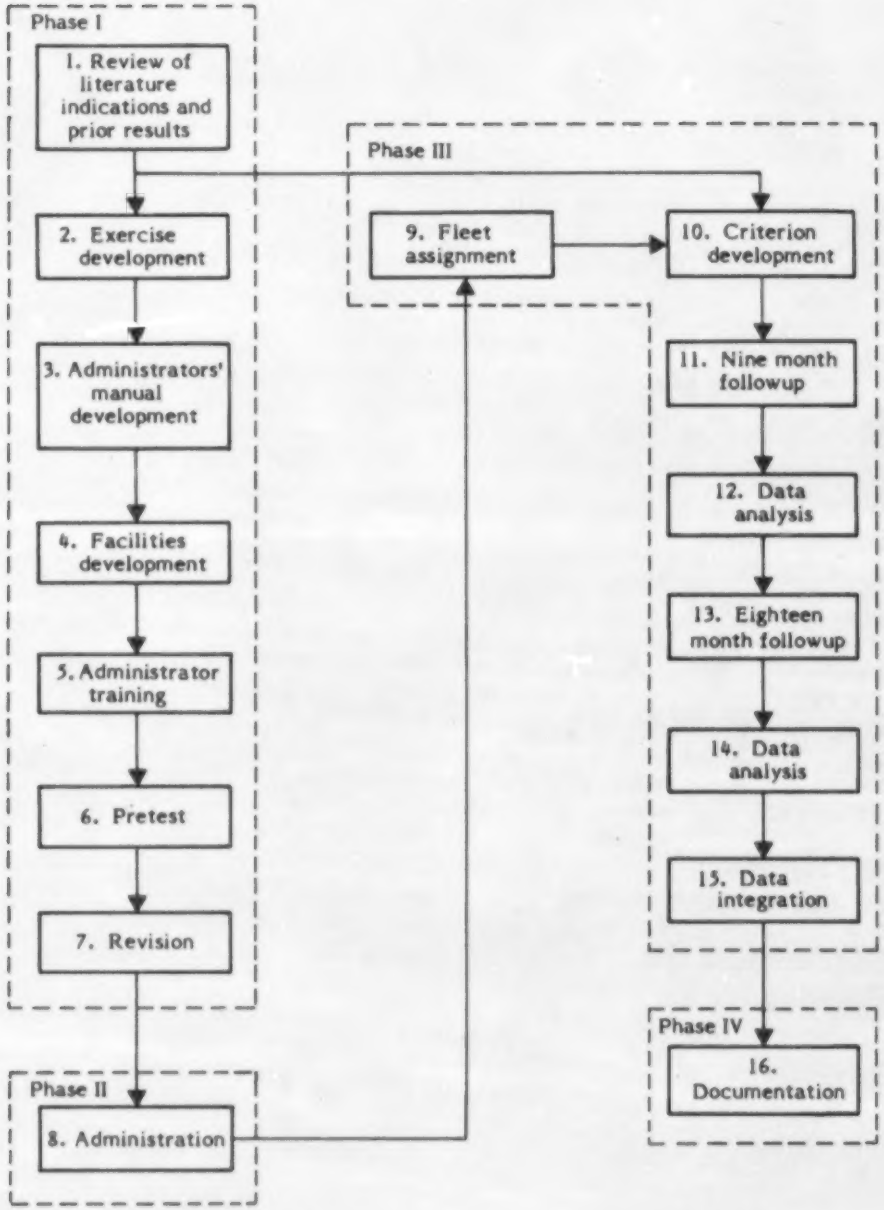

Figure 1. Program sequence.

Ten MTE exercises were created, which are briefly described below.

1. Computation and projection (CAP). In the training aspect of the CAP exercise, intercept course projection was taught. The recruits were shown how to read a simplified plot diagram of the positions of two ships, their headings and speeds, and to: (a) extrapolate the new position of each ship after 1 hour, and (b) evaluate the danger of collision (defined in the exercise as a ship separation of 1 mile or less). Simple addition,

subtraction, and ruler measurements were required to perform the work. After a training and a practice session, problems were administered in a group session. Three subscores (projection, collision identification, and course change direction) were derived, along with a total score.

2. Conceptual integration and application (CIA). In the CIA training session, the operation of an elementary, simulated electromechanical pumping system was explained and demonstrated. Potential malfunctions were diagnosed and the cause-effect relationships behind each diagnosis were explained. This ability to integrate facts and system relationships and derive a conclusion about the cause of a malfunction is commonly termed "trouble-shooting." After the training session and a set of practice problems, 10 problems involving system malfunction symptoms were group administered. The task was to state the cause of each malfunction. Two subscores (the number correct and the number incorrect) and the total score (right minus wrong) were derived.

3. Tool and object nomenclature, use, and recognition (TAO). The TAO exercise was used to measure the ability to learn the name and use of objects, which is requred at the entry level for most Navy ratings. During the training, 10 different Navy-oriented items were displayed and named and the use of each was described. The items were passed among the recruits who were encouraged to voice the name of each while they held and inspected them. This holding and voicing of the name was considered to constitute practice. The test situation was based on the recruit's ability to remember the names and uses of the objects. The evaluation was group-administered and yielded three subscores (names, use, and recall) and a total score, which was the sum of the three subscores.

4. Dual task (DT). The ability to share time between the performance of two or more different tasks is required for many Navy jobs. The two tasks used in the DT exercise were (a) simulated watchstanding and (b) fabricating a pipe assembly. In the watchstanding aspect, the recruit attended to a series of "alarm" lights and recorded their time of occurrence. The pipe assembly task included assembly of a set of pipes in accordance with a schematic diagram. During the training aspect of the exercise, the recruits were taught how to interpret and work with the schematic and how to make the required measurements. In the evaluation aspect, the recruits were required to fabricate the pipe assembly while simultaneously monitoring the "alarm lights" and entering their time of activation on a simulated log form. The test was group administered and yielded two subscores (pipe assembly and alarm recognition) and a total score.

5. Inspection/sort (SOR). In the SOR exercise, the recruits were shown "good" and "defective" items and were given practice in recognizing defects. The practice was monitored and feedback of results was provided. In the performance part of the exercise, the recruits were asked to detect imperfections or deviations from standards in five categories of objects. Nondeviant objects were placed in the appropriate bin and deviant objects were placed in a "reject" bin. Three subscores (speed, good items--accuracy, and bad items--accuracy) and an overall score were derived.

6. Record keeping (RK). In the RK exercise, the subjects were taught how to read displays of ship's speed and heading and how to log their values and "out of tolerance" conditions. Such an exercise measures general alertness, short term memory for numbers, and ability to evaluate information against a structured criterion. In the evaluation part of the exercise, 20 problems were presented. These were scored to yield five subscores (heading, heading out of tolerance, speed, speed out of tolerance, and time) and a total score.

7. Social interaction (SI). The SI exercise was used to measure subjects' work-group interactive tendencies and role assumption tendencies. There was no training for this exercise. From three to six recruits worked as a team to develop a method for folding a large tarpaulin into a form indicated on a schematic diagram. Four problems were included, each involving a planning and a plan-execution phase. Each team member was rated on cooperation, leadership, motivation, rule breaking, shirking, and interference. A total score was also derived.

8. Precision and planning (PP). In the PP exercise, the recruits were trained on how to work with an orthographic drawing to produce an object. After completing the training, the recruits were given another orthographic drawing and asked to fabricate the depicted object out of clay using a T-square and a knife. Successful completion of such an exercise involves planning an approach to work, careful measurement, precise hand-eye coordination, and three dimensional visualization. The evaluation situation was group administered and three subscores (dimensions, surface quality, and angles), plus a total score, were derived.

9. Relating diagrams with objects (ET). In the training for the ET exercise, symbols representing standard representations of electronic and electrical components were presented to the recruits. The components were placed on a table together with their associated symbols. The recruits walked around the table and studied the associations. They were allowed further time to study a card showing pictures of the objects and the associated symbol for each component, side-by-side. The recruits were next taught to fabricate a simple electrical assembly on the basis of a diagram showing interconnected symbols and were asked to construct it themselves. This assembly was checked and individual errors were corrected. In the evaluation aspect, five somewhat more complex assembly tasks were presented. The evaluation was group-administered to yield a score for each task along with a total score.

10. Level of aspiration (LOA). The LOA exercise provided a measure of need for achievement, tendency to readjust goals based on current performance, realism, and predisposition toward optimism or pessimism. The vehicle for these measurements was an exercise in which the recruit made an initial estimate of his expected performance on the basis of his own self appraisal and the administrator supplied norms and subsequent estimates of future performance with his own performance record on hand. Training was not incorporated into this exercise because predispositional constructs were involved and the recruit's typical approach to a task was desired. The results yielded measures of need for achievement, realism, optimism, pessimism, and a total score. The exercise was group-administered.

Administator's Manual and Facilities

For each MTE exercise, a detailed administrator's manual was developed that provided a full description of the procedures, including a standardized text for training, standardized instructions for administering tests, and full instructions for scoring and timing. Other sections of each administrator's manual covered such topics as: (1) the nature, scope, and purpose of the present program, (2) background information on the MTE approach, (3) samples of various session management forms, (4) the schedule to be followed, (5) privacy rights of the individuals assessed, and (6) special problems and how to manage them.

The battery was administered at the Naval Training Center (NTC), San Diego, in a suite of two large rooms. One room was used largely for training purposes; and the other for administering the evaluation aspects of the battery. Individual stations were provided

for each person being evaluated. Chalk-boards, projectors, and full training amenities were available for the purposes of the program.

Administrator Training, Pretest, and Revision

The administrator's manual was developed partially to assure full professionalism and standardization within the program. To further assure that these goals would be met, formal administrator training was implemented. This training involved 3 days and included a detailed formal review of and elaboration of the materials in the manual (the administrators studied the manual during the first 4 hours of the training).

The administrator training also included (1) a practice exercise administered by each administrator with a critique by the exercise developers, (2) a dry run of 1 day's operation, assessing recruits and employing the precise methods to be employed in conducting the program, and (3) a review and discussion of the "dry run" by the staff. The "dry run" indicated the need for a number of changes to the training and the evaluation materials. These changes were made before the exercises were administered.

Phase 2--Administration

Personnel

The MTEB was administered to 1034 male enlisted personnel at the Apprentice School, NTC, San Diego, during the period March 1978 through September 1978. These persons were in either the seaman, fireman, or airman ratings and were subsequently assigned to the fleet as GENDETs. All persons in the sample had been previously judged by customary methods to be ineligible for admission to an "A" school.

Two psychologists and one Navy chief petty officer administered the exercises. When the number of recruits available dropped to 30 or fewer per week, the exercises were administered by one psychologist and an assistant.

Initial Briefing of the Recruits

At the start of each administration period, the administrators met with the recruits for an introductory briefing. After the administrators introduced themselves and read a privacy act statement aloud, they asked the recruits if they had been tested before at the Center. Anyone answering affirmatively was dismissed. (Only a few such instances occurred). Also, any recruit with a watch before noon was identified, and his name transmitted to the appropriate authority so that a substitute watchstander could be appointed. Finally, the recruits were told that they would participate in a set of exercises that were important to the Navy and they were encouraged to try to do well on each.

When the administrators were speaking with the recruits, they were very friendly, using nontechnical terms and colloquialisms in order to promote understanding and maintain a hospitable climate. All important points were repeated with different emphasis to provide full understanding. The recruits were told to consider the "exercises" as training and not as tests. The recruits maintained general attention and interest throughout the session, and usually regarded the exercises as a welcome break in their apprentice training program. In a few cases, a recruit thanked the test administrators for the training given during the administration.

Training

Most kinds of testing are achievement-oriented; that is, test performance is based not only on the subject's ability, but also on his or her relevant experience. A major limitation of achievement-oriented testing, consequently, is that a person with ability in a particular field, but little or no training, may score poorly. The MTE exercises were designed to provide the training needed to bring out more fully the recruits' abilities. The types of training provided for each of the exercises are listed in Table 1 and defined below.

1. Tell. Tell the recruits what they are to do and how to do it. Provide all of the information needed to perform the exercise.

2. Question. Encourage questions and try to verify that the procedures are understood. Question each person to assure he or she understands the information presented.

3. Show. Show what is to be done. Solve sample problems showing clearly each step in the task. Use tools or make measurements in the manner that is most appropriate.

4. Practice. Have recruits practice task performance.

5. Feedback. Provide feedback during practice concerning what is correct or incorrect and make constructive suggestions.

6. Attention. Provide personal attention where problems exist.

Table 1

Exercise Administration Time and Type of Training Included

Exercise	Types of Training						Time (Min.)
	Tell	Question	Show	Practice	Feedback	Attention	
1. Computation and projection (CAP)	x	x	x	x	x	x	35
2. Conceptual integration and application (CIA)	x	x	x	x	x	x	50
3. Tool and object nomenclature, use, and recognition (TAO)	x	-	x	x	x	-	35
4. Dual task (DT)	x	x	x	x	x	-	30
5. Inspection/sort (SOR)	x	x	x	x	x	x	15
6. Record keeping (RK)	x	x	x	x	x	x	35
7. Social interaction (SI)	x	x	-	-	-	-	20
8. Precision and planning (PP)	x	x	x	x	x	x	30
9. Relating diagrams with objects (ET)	x	x	x	x	x	x	20
10. Level of aspiration (LOA)	x	-	-	-	-	-	10

Exercise Completion Time

The MTE approach to testing has been criticized because it requires more time to train and test than it takes to test. The 4- to 5-hour administration period is not excessive, however, when one considers the social and the financial costs of misclassification. The administration time for administering each exercise is included in Table 1.

Phase 3--Validation

To assess the MTEB's predictive validity, four criterion measures were developed, which are described below.

1. Supervisory Job Skill Technical Rating (SRAT). The job skill technical rating forms attempted to gain information about the recruit's ability to perform the technical aspects of his work. Three such forms were constructed--one each for those in the seamen, firemen, and airmen ratings. Each form contained a list of the tasks that an entry level person may perform, which were taken from NAVPERS 18068D, Section 1, Navy Enlisted Occupational Standards. The seaman job skill technical form completed by supervisors is shown in Appendix A. On a parallel form, each seaman rated himself on his performance on the same items included in the form completed by supervisors (RRAT). Scoring of each checklist was completed by calculating the average across items. Note that the result is an inverse scale with a lower score indicating superior performance.

2. Most Recent NAVPERS 792 Ratings. The NAVPERS 792 ratings, the official Navy enlisted performance evaluations, were also collected. The differences between the 792 and the SRAT ratings is that the 792 ratings represent the most recent "official" rating of enlisted personnel and may or may not have been made by the "immediate supervisor." They include ratings of the recruit's overall proficiency (SPROF), military behavior (SMILBEH), military appearance (SMILAPP), and adaptability (SADAPT). The ratings were obtained along a 9-point scale ranging from 1.0 to 4.0, with 4.0 indicating the best possible score. This scale is shown at the bottom of page A-2.

3. Supervisor Rating of Adjustment (SADJUST). Each recruit's adjustment to the Navy was rated on a 25-item checklist containing a set of descriptions assumed to reflect general satisfaction with Navy work and life. The items were drawn from prior factor analytic studies of Navy satisfaction (Siegel, Pfeiffer, & Braunstein, 1977). The immediate supervisor of each recruit completed the 25-item form using a 7-point category scale, where 1 = "Describes the man very well" and 7 = "Describes the man very poorly" (Appendix B). Like the job skill technical form, the adjustment form yielded an inverse scale (i.e., lower scores indicated superior adjustment).

These criterion measures were obtained approximately 9 months and 18 months after sample members began their first assignments. Their commanding officer (CO) was requested to have each recruit's immediate supervisor complete the job skill technical, overall, and adjustment rating forms and supply the most recent NAVPERS 792 ratings for each recruit. Table 2 presents the number of responses for each follow-up; and Table 3,

the number of seamen, firemen, and airmen cases available for analysis. Since, in some cases, all of the requested data were not received for a recruit, the number of cases involved in various analyses vary. In some of the correlational analyses, the N was further reduced because such analyses depend on matched pairs of data.

Table 2

Number of Responses for Each Follow-up

Group	First Follow-up	Second Follow-up
Seamen	275	189
Firemen	72	54
Airmen	84	62
	431	305

Table 3

Maximum Number of Cases Available for Analysis

Variable	Acronym	N			
		Firemen	Seamen	Airmen	Total[a]
MTE Variables—Total Score					
Computation and projection	CAP	179	613	241	1034
Conceptual Int. and Application	CIA	179	613	241	1034
Tool and object	TAO	179	613	241	1034
Dual task	DT	179	613	241	1034
Inspection/sort	SOR	179	613	241	1034
Record keeping	RK	179	613	241	1034
Social Interaction	SI	179	613	241	1034
Precision and planning	PP	96	216	112	425
Relating diagrams with objects	ET	97	216	112	426
Level of aspiration	LOA	179	613	241	1034
Total test score	TT	96	215	112	424
Criterion Ratings (1st Follow-up)					
Sup. job skill technical	SRAT1	69	258	78	405
Sup. overall proficiency	SPROF1	71	271	83	425
Sup. military behavior	SMILBEH1	71	275	84	430
Sup. military appearance	SMILAPP1	71	274	84	429
Sup. adaptability	SADAPT1	71	274	84	429
Sup. adjustment	SADJUST1	71	271	84	427
Recruit job skill technical	RRAT1	69	265	82	416
Criterion Ratings (2nd Follow-up)					
Sup. job skill technical	SRAT2	53	189	62	305
Sup. overall proficiency	SPROF2	53	182	61	297
Sup. military behavior	SMILBEH2	53	183	61	298
Sup. military appearance	SMILAPP2	53	183	61	298
Sup. adaptability	SADAPT2	53	183	61	298
Sup. adjustment	SADJUST2	54	188	62	305
Recruit job skill technical	RRAT2	53	186	60	300
ASVAB Subtest Scores					
General Information	GI	166	560	226	953
Numerical Operations	NO	166	560	226	953
Attention to Detail	AD	166	560	226	953
Word Knowledge	WK	166	560	226	953
Arithmetic Reasoning	AR	166	560	226	953
Space Perception	SP	166	559	225	951
Mathematics Knowledge	MK	166	539	225	951
Electronic Information	EI	166	559	225	951
Mechanical Comprehension	MC	166	559	225	951
General Science	GS	166	559	225	951
Shop Information	SI	166	559	225	951
Automotive Information	AI	166	559	225	951

[a]Where cases could not be identified, the sum of airmen, seamen, and firemen does not equal the total.

RESULTS

Evaluation of the findings involved analyses of the individual MTE exercises and of the MTE battery.

Mean and Standard Deviations of Each Exercise

Generally, the individual exercises evidenced adequate discriminating power. Table 4, which presents the mean and standard deviation for each MTE exercise, shows that individual exercise scores generally fell from the middle to the upper end of the range involved. The ET exercise, which showed the most depressed mean relative to the maximum possible score, was known at the outset to be difficult. It was included, in part, to obtain some discrimination at the upper end of the distribution of the low aptitude recruits. For the remaining exercises, the mean scores suggest that the recruits to whom the exercises were administered were provided with some success experience during the course of the exercises and, by implication, that the training sessions were meaningful. More importantly, the results suggest that, with the exception noted, the exercises were not "too difficult" for the low aptitude group.

Table 4

Exercise Means, Standard Deviations, and Maximum Possible Ranges

	Variable (Total Score)	Mean	Mean ÷ Total Possible (%)	Possible Range	Standard Deviation
1.	CAP	32.7	82	0 to 40	7.2
2.	CIA	5.5	71	-13 to 13	5.8
3.	TAO	24.0	80	0 to 30	5.1
4.	DT	19.5	78	0 to 25	5.9
5.	SOR	74.5	83	0 to 90	14.9
6.	RK	86.4	86	0 to 100	15.0
7.	SI	6.8	73	-15 to 15	3.2
8.	PP	59.9	83	0 to 72	13.4
9.	ET	11.1	28	0 to 40	11.1
10.	LOA	-.1	—	No theoretical range	4.4

Note. Ns are given in Table 3.

The standard deviations for the exercises, aside from that for the LOA exercise, ranged from 17 to 105 percent of the mean (RK and CIA exercises respectively) with the median percentage being 26.5 percent. These data suggest that the various exercises were able to differentiate among the recruits.

Score Distributions

Most of the score distributions were fairly normal but they tended to be negatively skewed (i.e., the scores tended to fall towards the higher end of the scale). However, in no case was the maximum or the minimum possible score attained by any recruit. The score distributions were all relatively smooth and continuous.

Intercorrelations Among Exercises

It is desirable that each MTE exercise be unique; if two exercises measure the same thing, one is not needed. Table 5 indicates that the intercorrelations between the various exercises were satisfactorily low. Of the 45 possible Pearson product-moment correlation coefficients, 30 fell below .20. The range was from -.05 (SOR and LOA) to .54 (RK and CAP). As might have been anticipated, the correlations between those for the LOA exercise and other exercises were lower than those among the other exercises. The LOA exercise is a motivational variable while the others (other than SI) were task-performance-oriented.

Table 5

Intercorrelations Among Total Test Scores on Miniature Training and Evaluation Exercises

MTE Scores	MTE Scores									
	CAP	CIA	TAO	DT	SOR	RK	SI	PP	ET	LOA
1. CAP	—	.40	.39	.24	.03	.54	.21	.16	.22	.09
2. CIA			.37	.15	.01	.35	.24	.13	.28	.08
3. TAO				.16	-.05	.37	.26	.18	.26	.04
4. DT					.10	.17	.11	.08	.12	.05
5. SOR						.05	-.01	.08	.15	-.05
6. RK							.18	.13	.22	.06
7. SI								.06	.24	.00
8. PP									.26	.07
9. ET										.14
10. LOA										—

Factor Analysis of MTEB

To develop additional insight into the MTEB's measurement attributes, all subscores and total scores from all exercises were factor analyzed by the principal components method with orthogonal rotation according to the varimax criterion. One additional measure, a measure of educational disadvantage (ED), was included in the analysis to assess the degree of support or orientation toward education received by the recruit from his home, his school, and his peer groups. This measure, which included 16 items, was administered to the entire sample (N = 1034) at the time of the MTEB administration for research purposes unrelated to MTE; it was included in this analysis to illustrate that the MTE predictions are not confounded by educational disadvantage.

Fourteen factors with eigenvalues greater than 1.00 were extracted and named. Table 6 presents these factors, their component exercise scores and subscores, loadings, eigenvalues, and percent of variance. As shown, the eight task-oriented MTE exercises

Table 6

Results of Factor Analysis of 48 Variables

	Emerging Factor	Component Exercise Total Score/Subscore	Loading	Eigenvalue	% Variance
1.	Record keeping[a]	Record keeping (RK):		9.77	20.4
		a. Total score	.95		
		b. Speed out of tolerance subscore	.76		
2.	Social facilitory	Social interaction (SI):		3.61	7.5
		a. Total score	.96		
		b. Cooperation subscore	.91		
3.	Tool and object[a]	Tool and object nomenclature, use, and recog. (TAO):		3.37	7.0
		a. Total score	.93		
		b. Name recognition subscore	.79		
4.	Diagram interpretation[a]	Relating diagrams with objects (ET):		2.63	5.5
		a. Test 1	.96		
		b. Test 3	.73		
5.	Precision and planning[a]	Precision and planning (PP):		2.48	5.2
		a. Total score	.96		
		b. Surface quality subscore	.88		
6.	Computation and projection	Computation and projection (CAP):		2.28	4.8
		a. Total score	.86		
		b. Projection subscore	.84		
7.	Conceptual integrative[a]	Conceptual integration and application (CIA):		2.03	4.2
		a. Total score	.93		
		b. Number correct	.87		
8.	Level of aspiration	Level of aspiration (LOA):		1.91	4.0
		a. Pessimism subscore	-.93		
		b. Total score	.83		
9.	Inspection/sort[a]	Inspection/sort (SOR):		1.68	3.5
		a. Speed subscore	.96		
		b. Total score	.93		
10.	Dual task[a]	Dual task (DT):		1.51	3.1
		a. Total score	.86		
		b. Pipe assembly subscore	.83		
11.	Realism/need for achievement	Level of aspiration (LOA):		1.35	2.8
		a. Realism subscore	.86		
		b. Need for achievement subscore	.56		
12.	Social interactive A	Social interaction (SI): Interference subscore	.80	1.16	2.4
13.	Social interactive B	Social interaction (SI): Rule-breaking subscore	.86	1.12	2.3
14.	Educational Disadvantage	Educational disadvantage measure	.84	1.02	2.1

[a]Factor emerging from task-oriented exercise.

yielded a single factor--factors 1, 3, 4, 5, 6, 7, 9, and 10 of Table 6. Since the total exercise score was the variable with the highest loading in six of these eight cases, it appears that the task-oriented MTE exercises are unique.

The social and motivational exercises, LOA and SI, fragmented into more than one factor. These factors were, however, distinct from the factors of the task-oriented exercises. LOA fragmented into factors 8 and 11, with factor 8 loading highest on the LOA pessimism subscore and total score; and factor 11, on the realism and need for achievement subscores. The SI exercise fragmented across one fairly strong factor (2) and two relatively weak factors (12 and 13). Factor 2 loaded highest on the SI total score and cooperation subscore; factor 12, on the SI interference subscore; and factor 13, on the rule breaking subscore. Interference involved placing objects, physical or mental, in the way of a quick and easy solution to a problem (e.g., causing conflict, arguing, showing distaste of an opinion through work or behavior), as well as chronically disrupting the performance of others attempting to perform their part of the task. Rule breaking involved accidental or intentional violation of the rules set down in the instructions to the exercise.

Educational disadvantage emerged as an individual, weak factor (14). Since the task-oriented exercises loaded very weakly, if at all, on this factor, some support is given to the contention that the MTE procedures minimize a recruit's educational disadvantage.

Discussion of Internal Characteristics

The analysis of the internal characteristics of the individual exercises within the MTE series suggests that each exercise provides measurement aspects that are not provided by the others. Moreover, on the basis of their standard deviations, the individual exercises seem to possess adequate discriminating power. In these regards, it is noted that the present sample of recruits was drawn from the lower range of the total recruit population. All were judged, on the basis of the usual classification methods, as ineligible for entry into a Navy "A" school. If a more heterogenous sample was involved, greater standard deviations and higher mean scores would undoubtedly be attained. On the basis of these results, there seems to be little to disqualify the MTE concept or the exercises used in the present series.

Predictive Validity Analyses

As indicated previously, to assess the MTEB's predictive validity, criterion measures were developed and data were collected for the 1034 sample members 9 months and 18 months after they had been assigned to the fleet. At 9 months (first follow-up), data were received for 431 recruits, compared to 305 at 18 months (second follow-up).

The primary criterion was considered to be the supervisory rating of the recruits' performance on the tasks included in the pertinent job skill technical form (SRAT), because the MTE exercises were task-oriented. Individual proficiency ratings were obtained on about 30 individual job tasks and an average score was derived for each recruit. The supervisory rating of overall performance (SPROF) represented a secondary criterion, with results for seamen assuming greater importance than those for firemen and airmen because the exercises were designed with deck (seamen) specialties in mind. Other rating data of interest, but of considerably less relevance, were the supervisor ratings of military behavior (SMILBEH), military appearance (SMILAPP), adaptability (SADAPT), and adjustment (SADJUST).

Although supervisory ratings are subject to a wide variety of error sources, they represent about the only practical criterion approach in many situations. Other types of

follow-up criteria, although possible, are very costly when a sample is dispersed throughout the world, as are Navy personnel. Although training programs can be instituted to control for or correct common rater errors, such programs are often impractical in the Navy situation. A further source of error in the present criterion data collection approach lies in the lack of control over the time and effort a supervisor devotes to the rating process. Some supervisors may take the rating task very conscientiously, while others may have a more carefree attitude toward it. Nevertheless, in the day-to-day work situation, a journeyman's rate of progress is probably best known by his immediate supervisor.

Response to Follow-ups

For the total of 1034 enlisted personnel in the original sample, addresses were available for 929 in the first follow-up and 838 in the second. As indicated previously, criterion data returns were received for 431 persons (46%) in the first follow-up and 305 (36%) in the second. About 33 percent of all Navy recruits are expected to leave the Navy for one reason or another sometime during their first 18 months of service. When one adds to this percentage a number of wrong addresses due to transfers and a number of "no responses," a sample shrinkage is expectable.

There is reason to believe that those who possess lower aptitude, even within the low aptitude range of this sample, tend toward early separation. If this conjecture is true, then the separations would not be randomly distributed and there would be a truncation of the response distribution that would result in a more homogeneous set of results. These effects would ordinarily serve to decrease predictive correlation coefficients.

Product-moment Correlations

Criterion intercorrelations are presented in Table 7. Product-moment correlations between MTE predictors and the criteria for the first and the second follow-ups are presented in Table 8 and discussed below.

Seamen. For the SRAT criterion, all but one of the individual exercise-criterion correlation coefficients were in the expected direction. For the first follow-up, the correlations between SRAT and the MTE exercise scores ranged from .02 (CIA) to -.23 (PP), with the median being -.17. All of the SRAT correlations were in the expected direction and six of them were statistically significant at or below the .05 level of confidence. For the second follow-up, some shifts in magnitude were evidenced. The median correlation between SRAT and the MTE exercises was -.16, and the correlation of SRAT with the total MTEB score was -.30.

The relationships between the individual exercises and the overall rating provided by the supervisors (SPROF) were weaker. This was to be anticipated because overall ratings are typically subject to greater error than are individual task or trait ratings. Nonetheless, the correlations between the total MTFB score and SPROF for the first and second follow-ups were .19 and .33 respectively.

For the remaining four criteria, regardless of follow-up, most of the individual exercise-criterion correlations were in the expected directions and of low magnitude. This trend appears to be appropriate because the MTE exercises were not designed to predict the behavior implied by these criteria.

Firemen. For firemen, the zero-order correlations between SRAT and the individual MTE exercise scores were, by and large, in the expected direction and of moderate magnitude. The median correlations were -.22 for the first follow-up and -.13 for the

Table 7

Intercorrelations Among Criteria

Criterion	SRAT	SPROF	SMILBEH	SMILAPP	SADAPT	SADJUST
Seamen						
First Follow-up:						
SRAT	---	-.69*	-.59*	-.58*	-.60*	.72*
SPROF	(246)		.84*	.75*	.86*	-.81*
SMILBEH	(250)	(268)		.72*	.86*	-.77*
SMILAPP	(249)	(268)	(271)		.70*	-.70*
SADAPT	(249)	(268)	(271)	(271)		-.76*
SADJUST	(248)	(259)	(263)	(262)	(262)	---
Second Follow-up:						
SRAT	--	-.49*	-.41*	-.37*	-.55*	.64*
SPROF	(181)		.63*	.55*	.69*	-.54*
SMILBEH	(182)	(182)		.73*	.82*	-.55*
SMILAPP	(182)	(182)	(183)		.73*	-.41*
SADAPT	(182)	(182)	(183)	(183)		-.68*
SADJUST	(184)	(178)	(179)	(179)	(179)	---
Firemen						
First Follow-up:						
SRAT	---	-.73*	-.56*	-.55*	-.68*	.63*
SPROF	(076)		.75*	.62*	.83*	-.70*
SMILBEH	(076)	(078)		.71*	.81*	-.76*
SMILAPP	(076)	(078)	(078)		.67*	-.63*
SADAPT	(076)	(078)	(078)	(078)		-.78*
SADJUST	(076)	(078)	(078)	(078)	(078)	---
Second Follow-up:						
SRAT	---	-.84*	-.65*	-.58*	-.48*	.80*
SPROF	(052)		.69*	.67*	.48*	-.78*
SMILBEH	(052)	(053)		.63*	.64*	-.68*
SMILAPP	(052)	(053)	(053)		.41*	-.59*
SADAPT	(052)	(053)	(053)	(053)		-.59*
SADJUST	(053)	(053)	(053)	(053)	(053)	---
Airmen						
First Follow-up:						
SRAT	---	-.65*	-.43*	-.63*	-.71*	.70*
SPROF	(072)		.77*	.78*	.83*	-.70*
SMILBEH	(073)	(079)		.80*	.69*	-.60*
SMILAPP	(073)	(079)	(080)		.80*	-.65*
SADAPT	(073)	(079)	(080)	(080)		-.72*
SADJUST	(074)	(078)	(079)	(079)	(079)	---
Second Follow-up:						
SRAT	---	-.52*	-.42*	-.30*	-.61*	.77*
SPROF	(061)		.68*	.65*	.73*	-.74*
SMILBEH	(061)	(061)		.61*	.88*	-.72*
SMILAPP	(061)	(061)	(061)		.68*	-.66*
SADAPT	(061)	(061)	(061)	(061)		-.80*
SADJUST	(062)	(061)	(061)	(061)	(061)	---

Note. Sample sizes are in parentheses.

*$p \leq .05$.

Table 8

Correlations Between MTE Predictors and Criteria

MTE Exercises (Total Score)	Criterion					
	SRAT	SPROF	SMILBEH	SMILAPP	SADAPT	SADJUST
Seamen						
First Follow-up:						
CAP	-.19*	.08	.04	.05	.05	-.06
CIA	.02	-.02	-.07	-.05	-.04	-.03
TAO	-.18*	.10*	.06	.06	.12*	-.11*
DT	-.14*	.02	.01	.06	.06	-.03
SOR	-.09	-.01	.03	.08	.08	-.13*
RK	-.09	.06	.00	.03	.02	-.02
SI	-.20*	.05	.04	.07	.04	-.11*
PP	-.23*	.27*	.21*	.18*	.26*	-.21*
ET	-.11	.17*	.09	.12	.14	-.09
LOA	-.03	.07	.02	.04	.02	-.03
TT	-.37*	.19*	.12	.16	.18*	-.19*
Second Follow-up:						
CAP	-.27*	.20*	.10	.08	.17*	-.18*
CIA	-.20*	.17*	.18*	.14*	.15*	-.03
TA	-.18*	.00	-.03	-.02	-.01	.06
DT	-.14*	.10	.12	.13*	.09	-.04
SOR	-.02	-.02	.02	.05	.05	-.04
RK	-.22*	.17*	.15*	.11	.22*	-.14*
SI	.00	-.01	.01	.05	.02	.11
PP	-.32*	.30*	.32*	.11	.31*	-.32*
ET	-.09	.14	.19	.12	.19	-.22*
LOA	-.05	.05	.09	.11	.10	-.08
TT	-.30*	.33*	.27*	.16	.25*	-.20*
Firemen						
First Follow-up:						
CAP	-.30*	.35*	.12	.06	.35*	-.16
CIA	-.25*	.41*	.19	.18	.37*	-.19
TAO	-.34*	.33*	.24*	.03	.30*	-.15
DT	-.24*	.17	.07	.15	.21*	-.25*
SOR	.09	-.19	-.01	.10	-.01	-.04
RK	-.40*	.32*	.18	.03	.29*	-.09
SI	-.04	.10	-.04	-.03	.07	-.11
PP	.10	.21	.20	.40*	.14	-.04
ET	-.15	.32*	-.06	.17	.23	-.03
LOA	.20*	-.12	-.12	-.20*	-.03	.14
TT	-.34*	.35*	-.02	.04	.30*	-.11

Note. Ns are given in Table 3. Since the SRAT instrument was inversely scaled, all correlations associated with it were expected to be negative.

*$p \leq .05$.

Table 8 (Continued)

MTE Exercises (Total Score)	Criterion					
	SRAT	SPROF	SMILBEH	SMILAPP	SADAPT	SADJUST
Firemen (Continued)						
Second Follow-up:						
CAP	-.18	.19	.02	-.07	.00	-.15
CIA	-.11	.25*	.15	.11	.05	-.08
TAO	-.32*	.40*	.20	.10	.21	-.26*
DT	-.22*	.25*	.16	-.04	.07	-.24*
SOR	-.06	.00	.13	.08	.08	.01
RK	-.14	.21	.04	.06	.02	-.04
SI	-.12	.27*	.06	.00	.02	-.02
PP	-.17	-.02	-.13	-.14	-.25	-.13
ET	-.02	.13	.06	-.13	-.41*	.22
LOA	.10	-.17	-.21	.14	-.08	.12
TT	-.42*	.43*	.20	.15	.05	-.17
Airmen						
First Follow-up:						
CAP	.19	-.05	-.04	-.02	-.03	.24*
CIA	.19*	.06	-.00	-.01	.08	.04
TAO	-.17	.18*	.12	.10	.23*	-.06
DT	-.08	.09	.12	-.03	.13	-.07
SOR	-.15	.17	.26*	.23*	.15	-.12
RK	-.07	.20*	.19*	.11	.20*	-.07
SI	.03	.05	-.02	.04	-.00	-.01
PP	.01	.02	.10	.21	.07	.05
ET	.09	-.17	-.19	-.07	-.14	.25
LOA	.01	-.05	.07	-.02	.06	.07
TT	-.04	.01	-.01	.19	.14	-.03
Second Follow-up:						
CAP	.06	.13	-.09	.04	-.15	.17
CIA	.06	.03	.08	-.08	-.03	.12
TAO	-.22*	.27*	.14	-.01	.20	-.21
DT	.10	.01	-.17	-.15	-.23*	.20
SOR	-.25*	.13	.09	.07	.16	-.21
RK	.01	.02	-.10	-.05	-.05	.13
SI	-.02	-.29*	-.09	-.28*	-.05	.25*
PP	-.14	.01	-.10	-.13	-.05	.08
ET	-.19	-.05	-.01	-.16	.05	-.02
LOA	.00	-.05	-.22*	-.23*	-.19	.08
TT	-.31*	.12	.09	-.18	.15	-.14

Note. Ns are given in Table 3. Since the SRAT instrument was inversely scaled, all correlations associated with it were expected to be negative.

*$p \leq .05$.

second. The correlation between SRAT and the total MTEB score for the first and second follow-ups were -.34 and -.42 respectively, both of which are statistically significant.

In the first follow-up, four MTE exercises were significantly correlated with both SRAT and SPROF. SPROF coefficients ranged from -.19 to .41 for the first follow-up and from .17 to .40 for the second. The median correlation between the MTE exercises and SPROF for the first and second follow-ups were .25 and .20 respectively. For both follow-ups, the total score-SPROF correlations for firemen (.35 and .43) were greater than those for seamen (.19 and .33).

The correlations between the MTE exercises and the four nontechnical criteria were low and in the expected direction.

Airmen. For the airmen, correlations between the individual MTE exercises and SRAT and SPROF tended to be lower than those for the seamen and the firemen. The number of significant correlations of SRAT and SPROF with the MTE scores fell to seven for both follow-ups, and two of these were not in the expected direction.

Discussion

The full impact of these relationships can be understood only in the perspective of the criteria involved. One measure of criterion reliability is the correlations between the supervisors' judgments on the first and second follow-ups. These correlations, which are presented in Table 9, range was from .26 to .58, with a median of .40. If these values represent the true reliability of the criteria for the various groups, it appears that they are unacceptably low. Criterion reliability, in the general case, may be assumed to represent the upper level of possible validity. From this perspective, the zero order relationships assume increased impressiveness.

Table 9

Correlations Between Criterion Ratings Received in the Two Follow-ups

	Group			
Criterion	Seamen	Firemen	Airmen	Total
SRAT	.44	.48	.40	.43
SPROF	.39	.51	.44	.40
SMILBEH	.40	.31	.31	.36
SMILAPP	.26	.29	.51	.29
SADAPT	.43	.30	.33	.37
SADJUST	.47	.52	.58	.49

The restriction of range effects, mentioned earlier, also affect the predictive correlation coefficients. Accordingly, it is believed that the zero order relationships are depressed artificially and represent lower limits. Nonetheless, of the 22 possible correlations with SRAT for each group, 13 (59%) were statistically significant for seamen; 10 (55%), for firemen; and 4 (18%), for airmen. For SPROF, the numbers of significant correlations for seamen, firemen, and airmen were 9 (41%), 11 (50%), and 4 (18%), respectively. Moreover, for the seaman and the fireman groups, all the correlations of total score with SRAT and SPROF were statistically significant and of moderate magnitude.

Correction for Criterion Unreliability

To assess the effects of criterion unreliability on the zero-order relationships, the correction method given by Guilford and Fruchter (1978) was applied to the zero order relationships for the SRAT and the SPROF criteria. According to Guilford and Fruchter, such a correction is quite acceptable because one is interested in what the "true" correlation would be if one had a reliable criterion. On the other hand, corrections that manipulate the predictor (e.g., reliability of the exercises) are less acceptable because the predictor will not change in actual application. The "corrected" correlations, given in Table 10, have acceptable magnitudes in most cases. Further, except for the first follow-up for airmen, the correlations with total score are impressive.

Multiple Prediction

To investigate further the MTEB's power to predict the various criteria, a set of multiple correlation coefficients was computed. In these analyses, individual exercises were successively added until no more statistically significant variance could be accounted for by the inclusion of additional predictors. The results of these analyses are presented in Table 11. The number in parentheses that follows each multiple correlation coefficient indicates the number of predictors that entered into the calculation.

For the technical performance criteria, SRAT and SPROF, the multiple correlation coefficients were of reasonable magnitude, with 8 of 12 being statistically significant. Even for airmen, these multiple correlation coefficients assumed moderate values. These findings lend additional credence to the prior discussion, which supported the value of the present set of MTE exercises and the assessment concept involved.

Comparison of the MTEB with the ASVAB

The ASVAB, a battery of 12 paper-and-pencil tests administered at the recruiting station or soon after a recruit enters military service, is the primary psychometric instrument employed by the military services for selection/classification. Accordingly, since ASVAB scores were available for sample members (Table 3), it seemed proper to compare the predictive power of the ASVAB with that of the MTEB. This comparison provides additional insight into the value of the MTEB and is not intended as a critical analysis of the ASVAB itself.

The product-moment correlations between the ASVAB test and the follow-up criteria are presented in Table 12. When interpreting these correlations, the earlier comments about follow-up criteria still apply.

As for the MTE exercises, the most salient criteria are SRAT and SPROF, with SRAT assuming primacy. By way of overview, note that the uncorrected zero-order correlation coefficients relating the ASVAB tests to both SRAT and SPROF tended to be quite low. This finding was anticipated, since the conjecture that paper-and-pencil tests are not entirely appropriate for the people represented by our sample formed a partial basis for the present study.

The results in Table 12 and in Table 8 allow direct comparison of the zero order predictor-criterion relationships for the two batteries. For the first follow-up, and ignoring the total score predictor, the individual MTE exercises produced 22 statistically significant correlations with SRAT and SPROF out of a possible 60 (37%). The ASVAB tests produced 8 out of a possible 72 (11%). For the second follow-up, the comparable data were 20 (33%) for MTE and 11 (15%) for ASVAB. The differences for both follow-ups are statistically significant at or below the .05 level of confidence.

Table 10

MTE Exercise Validity Coefficients Corrected for Criterion Unreliability

MTE Exercises (Total Score)	First Follow-up SRAT	First Follow-up SPROF	Second Follow-up SRAT	Second Follow-up SPROF
Seamen				
CAP	-.29	.13	-.41	.32
CIA	.03	-.03	-.30	.27
TAO	-.27	.16	-.27	.00
DT	-.21	.03	-.21	.16
SOR	-.14	-.02	-.03	-.03
RK	-.14	.10	-.33	.27
SI	-.30	.08	.00	-.02
PP	-.35	.43	-.48	.48
ET	-.17	.27	-.14	.22
LOA	-.05	.11	-.08	.08
TT	-.56	.30	-.45	.53
Firemen				
CAP	-.43	.49	-.26	.27
CIA	-.36	.57	-.16	.35
TAO	-.49	.46	-.46	.56
DT	-.35	.24	-.32	.35
SOR	.13	-.27	-.09	.00
RK	-.58	.45	-.20	.29
SI	-.06	.14	-.17	.38
PP	.14	.29	-.25	-.03
ET	-.22	.45	-.03	.18
LOA	.29	-.17	.14	-.24
TT	-.49	.49	-.61	.60
Airmen				
CAP	.30	-.08	.09	.20
CIA	.30	.09	.09	.05
TAO	-.27	.27	-.35	.41
DT	-.13	.14	.16	.02
SOR	-.24	.26	-.40	.20
RK	.11	.30	-.02	.03
SI	.05	.08	-.03	-.44
PP	.02	.03	-.22	.02
ET	.14	-.26	-.30	-.08
LOA	.02	-.08	.00	-.08
TT	-.06	.02	-.49	.18

Table 11

Multiple Correlations Between Criteria and MTE Predictors

	First Follow-up			Second Follow-up		
Criteria	Seamen	Firemen	Airmen	Seamen	Firemen	Airmen
SRAT	.36 (08)*	.55 (10)*	.41 (10)	.41 (10)*	.42 (10)	.45 (10)
SPROF	.30 (10)*	.53 (08)*	.49 (10)*	.39 (10)*	.50 (09)	.50 (09)*
SMILBEH	.24 (09)	.32 (07)	.57 (10)*	.42 (10)*	.42 (09)	.43 (10)
SMILAPP	.24 (08)	.42 (10)	.42 (08)	.26 (10)	.36 (10)	.45 (09)
SADAPT	.40 (07)*	.42 (08)	.50 (09)*	.42 (10)*	.65 (10)*	.47 (10)
SADJUST	.27 (08)*	.31 (09)	.46 (09)*	.46 (09)*	.50 (10)	.58 (10)*

Note. Number of MTE predictors is shown in parentheses. Ns are given in Table 3.

*$p \leq .05$.

Table 12

Correlation Coefficients Between ASVAB Predictors and Criteria

Group ASVAB (Predictors)	Criteria					
	SRAT	SPROF	SMILBEH	SMILAPP	SADAPT	SADJUST
Seamen						
First Follow-up:						
1. GI	-.03	-.02	-.05	-.06	-.02	.02
2. NO	-.01	.04	-.02	.01	.03	.01
3. AD	-.14*	.00	-.02	.07	-.01	-.01
4. WK	.05	-.09	-.15*	-.10*	-.10*	.14*
5. AR	-.11*	.03	-.05	-.04	.03	.02
6. SP	.04	-.08	-.12*	-.08	-.07	.05
7. MK	-.02	.04	-.02	.01	.07	-.03
8. EI	-.01	-.02	-.05	-.02	.00	.05
9. MC	-.03	-.02	-.09	-.07	-.04	.04
10. GS	.01	-.03	-.06	-.07	-.01	.06
11. SI	-.00	-.04	-.08	-.07	-.01	.04
12. AI	-.05	.01	-.03	-.01	.04	.01
Second Follow-up:						
1. GI	.02	-.09	-.10	-.11	-.17*	.19*
2. NO	-.04	-.03	-.02	-.04	.02	.03
3. AD	-.12	.04	-.06	-.01	.01	.01
4. WK	.13*	-.17*	-.24*	-.16*	-.26*	.28*
5. AR	-.02	.01	.01	-.04	.02	.08
6. SP	-.02	-.09	-.12	-.13*	-.07	.07
7. MK	.00	-.09	-.03	-.06	.00	.05
8. EI	-.04	-.04	-.14*	-.07	-.10	.16*
9. MC	-.03	-.06	-.07	-.07	-.05	.06
10. GS	-.00	-.07	-.11	-.10	-.08	.12
11. SI	-.04	-.04	-.10	-.08	-.05	.14*
12. AI	-.02	-.04	-.12	-.10	-.09	.18*
Firemen						
First Follow-up:						
1. GI	.04	.05	.08	-.08	.05	.10
2. NO	-.27*	.05	.05	-.04	.09	-.18
3. AD	-.10	.26*	.20*	.02	.34*	-.18
4. WK	-.05	.03	.10	-.15	.02	.08
5. AR	-.11	.12*	.04	.01	.09	-.01
6. SP	-.09	.13	.16	.22*	.15	-.13
7. MK	-.12	.05	.13	.08	.13	.03
8. EI	-.16	.15	.22*	-.06	.11	-.04
9. MC	.02	.18	.12	.02	.12	-.01
10. GS	-.01	.01	-.02	-.12	-.08	.09
11. SI	-.21*	.20*	.04	.01	.07	-.01
12. AI	-.09	.11	.16	-.03	.09	.10

Note. Ns are given in Table 3.

*$p \leq .05$.

Table 12 (Continued)

Group ASVAB (Predictors)	Criteria					
	SRAT	SPROF	SMILBEH	SMILAPP	SADAPT	SADJUST
Firemen (Continued)						
Second Follow-up:						
1. GI	.07	.12	-.07	-.01	-.09	.03
2. NO	-.15	.23*	.10	.02	.23*	-.15
3. AD	-.14	.21	.09	.29*	.21	-.09
4. WK	.18	-.05	-.22	-.29*	.02	.16
5. AR	-.18	.28*	-.03	.10	.01	-.17
6. SP	.10	.12	.18	-.01	.11	.04
7. MK	.11	.15	-.09	-.18	-.04	-.01
8. EI	.11	.05	.00	-.05	.15	.19
9. MC	-.14	.28*	.16	.26*	.15	-.02
10. GS	.11	.05	-.25*	-.13	-.07	.17
11. SI	-.11	.28*	.13	.07	.07	-.02
12. AI	.10	.03	-.03	-.21	-.09	.22
Airmen						
First Follow-up:						
1. GI	.18	-.11	-.08	-.02	-.09	.19*
2. NO	-.06	.06	.02	.12	.06	.04
3. AD	-.07	.07	.12	.17	.11	-.12
4. WK	.09	-.20*	-.19*	-.19*	.24*	.26*
5. AR	-.02	.01	.04	.07	-.07	.02
6. SP	-.05	-.02	-.05	-.00	.00	.06
7. MK	-.04	-.10	-.18	-.05	-.05	.11
8. EI	-.02	.02	-.08	-.09	.06	.04
9. MC	.00	-.15	-.21*	-.25*	-.13	.21*
10. GS	.12	-.15	-.19*	-.21*	-.13	.25*
11. SI	-.04	.02	-.04	-.11	.06	.13
12. AI	.00	-.06	-.11	-.17	.04	.09
Second Follow-up:						
1. GI	.05	.07	-.05	.06	-.05	.02
2. NO	-.24*	.27*	.15	.19	.22*	-.12
3. AD	-.15	.29*	.19	.14	.17	-.09
4. WK	.40*	-.19	-.22*	-.27*	-.26*	.31*
5. AR	-.13	.09	-.03	.14	-.06	.08
6. SP	-.21*	.12	.10	.02	.06	-.03
7. MK	-.04	.06	-.05	.11	-.02	.04
8. EI	.12	-.10	-.14	-.14	-.16	.16
9. MC	.15	-.19	-.26*	-.25*	-.29*	.27*
10. GS	.09	-.05	-.15	-.09	-.14	.15
11. SI	.08	.06	-.03	-.05	-.09	.10
12. AI	.01	-.01	-.00	-.03	-.00	.03

Note. Ns are given in Table 3.

*$p \leq .05$.

Table 13, which presents the median and best uncorrected correlations MTEB and ASVAB with the follow-up criteria, shows that five of the six SRAT median coefficients favored the MTEB over the ASVAB. From the point of view of the "best" correlations between the two predictor sets and the SRAT criterion, all six correlation coefficients favored the MTEB. However, there was almost a tie for the airman group in the second follow-up. For the SPROF criterion, five of the six medians, and five of six "best" comparisons, favored the MTEB.

Table 13

Median (M) and Best (B) Uncorrected Correlation for MTEB and ASVAB for SRAT and SPROF Criteria

	First Follow-up				Second Follow-up			
	SRAT[a]		SPROF		SRAT[a]		SPROF	
	M	B	M	B	M	B	M	B
Seamen								
ASVAB	-.02	-.14	-.02	.04	-.02	-.12	-.04	.04
MTEB	-.17	-.23	.08	.27	-.16	-.32	.12	.30
Firemen								
ASVAB	-.10	-.27	.12	.26	.10	-.18	.14	.28
MTEB	-.22	-.40	.27	.41	-.13	-.32	.20	.40
Airmen								
ASVAB	-.03	-.07	-.02	.07	.03	-.24	.06	.29
MTEB	.01	-.17	.07	.20	-.01	-.25	.02	.27

[a]The SRAT instrument was inversely scaled and all correlations associated with it were expected to be negative.

The comparisons of ASVAB and MTEB suggest that the MTEB has superior predictive power--particularly for seamen and firemen. In the case of airmen, the two batteries appear to possess about equal predictive power.

Of course, the present findings are derived from a sample of low aptitude recruits, and a generalization of these findings to a moderate or a high aptitude group is not warranted. Because the MTE concept was proposed as an approach to evaluating the potential of low aptitude personnel, the comparative result may have been anticipated. However, since the ASVAB is addressed to a broader spectrum, it may be more valuable when broad spectrum assessment, or the assessment of higher aptitude personnel, is involved. Also, it should be noted that ASVAB subtest scores are not used singly. Rather, scores on two or more tests are combined for predictive purposes. There is reason to believe that such combinations may increase the predictive power of the ASVAB.

The zero-order relationships among the ASVAB scores and the criteria were corrected for criterion unreliability in the same manner as described previously for the

MTEB. The results are shown in Table 14. While the corrections more closely represent the "true" correlation between the ASVAB scores of the sample and the various criteria, they do not alter the interpretations presented earlier.

Comparison of Multiple Relationships

Further comparisons between the ASVAB and the MTEB may be made on the basis of multiple correlation coefficients. These yield the maximum covariance between predictors and a criterion when the various subtests are optimally used in combination. Table 15, which presents the multiple correlations between the ASVAB and the various criteria, shows that the multiple correlations tended to be higher for the second follow-up than for the first. Two of the 12 (17%) correlations for SRAT and SPROF were statistically significant, compared with 8 of 12 (67%) for MTEB (Table 11). Comparisons of SRAT and SPROF criteria in Tables 15 and 11 indicate that: (1) for the first follow-up, the MTEB was superior to ASVAB regardless of group, and (2) for the second follow-up, the MTEB was superior to ASVAB for the seaman group but about equivalent to ASVAB for the fireman and airman groups.

What the MTE Approach Adds

To determine how much predictive variance is added by the MTE exercises to that already available from the ASVAB, hierarchical multiple correlation coefficients were calculated according to the following logic:

1. Those ASVAB tests showing statistically significant zero order correlation coefficients with the SRAT criterion were added one at a time to the predictive equation.

2. The MTE exercises that evidenced statistically significant zero order relationships with the criterion were added.

The percentage of predictable variance added by the MTE approach over that provided by the statistically significant ASVAB predictors was then calculated.

For the first follow-up, the analysis was completed for seamen only and the number of cases was 258. For the second follow-up, the total of 305 seamen, airmen, and firemen with supervisory evaluations was included. Table 16, which presents the results of these analyses, shows that, for both follow-ups, there was considerable advantage to using the MTEB in conjunction with the ASVAB. These findings, which support the use of the MTEB in conjunction with ASVAB, reinforce prior contentions favoring the utility of the MTE exercises. Also, they confirm the findings of Cory (1982), who found advantage for the use of a bivariate model (ASVAB plus MTE exercises) in Navy selection/classification.

Table 14

ASVAB Validity Coefficients Corrected for Criterion Unreliability

ASVAB Predictor	First Follow-up		Second Follow-up	
	SRAT	SPROF	SRAT	SPROF
Seamen				
1. GI	-.05	-.03	.03	.14
2. NO	-.02	.06	-.06	.05
3. AD	-.21	.00	-.18	.06
4. WK	.08	-.14	.20	.27
5. AR	-.17	.05	-.03	.02
6. SP	.06	-.13	-.03	.14
7. MK	-.03	.06	.00	.14
8. EI	-.02	-.03	-.06	.06
9. MC	-.05	-.03	-.05	.10
10. GS	.02	-.05	-.00	.11
11. SI	-.00	-.06	-.06	.06
12. AI	-.08	.02	-.03	.06
Firemen				
1. GI	.06	.07	.10	.17
2. NO	-.39	.07	-.22	.32
3. AD	-.14	.36	-.20	.29
4. WK	-.07	.04	.26	.07
5. AR	-.16	.17	-.26	.39
6. SP	-.13	.18	.14	.17
7. MK	-.17	.07	.16	.21
8. EI	-.23	.21	.16	.07
9. MC	.03	.25	-.20	.39
10. GS	-.01	.01	.16	.07
11. SI	-.30	.28	-.16	.39
12. AI	-.13	.15	.14	.04
Airmen				
1. GI	.28	-.17	.08	.11
2. NO	-.09	.09	-.38	.41
3. AD	-.11	.11	-.24	.44
4. WK	.14	-.30	.63	-.29
5. AR	-.03	.02	-.21	.14
6. SP	-.08	-.03	-.33	.18
7. MK	-.06	-.15	-.06	.09
8. EI	-.03	.03	.19	-.15
9. MC	.00	-.23	.24	-.29
10. GS	.19	-.23	.14	-.08
11. SI	-.06	.03	.13	.09
12. AI	.00	-.09	.02	-.02

Table 15

Multiple Correlations Between ASVAB and Follow-up Criteria

	First Follow-up			Second Follow-up		
Criteria	Seamen	Firemen	Airmen	Seamen	Firemen	Airmen
SRAT	.24 (12)	.43 (11)	.29 (12)	.23 (10)	.48 (10)	.66 (12)*
SPROF	.17 (09)	.34 (12)	.39 (11)	.24 (11)	.52 (12)	.56 (12)*
SMILBEH	.20 (12)	.38 (12)	.43 (12)	.30 (11)	.52 (12)	.52 (10)*
SMILAPP	.20 (12)	.35 (11)	.42 (11)	.21 (09)	.67 (12)*	.57 (12)*
SADAPT	.22 (11)	.41 (12)	.47 (11)	.35 (11)*	.48 (12)	.58 (10)*
SADJUST	.19 (11)	.35 (10)	.44 (12)	.34 (12)*	.43 (12)	.50 (11)

Note. Number of included variables is shown in parentheses. Ns are given in Table 3.

*$p < .05$.

Table 16

Multiple Correlation (R) and Percentage of Contribution to Total Predictable Variance (PC) for SRAT Criterion

	ASVAB Only		ASVAB Plus MTEB	
Follow-up	R	PC	R	PC
First	.20	40	.32	60
Second	.24	35	.41	65

DISCUSSION

Feasibility of the MTE Approach

While the MTE exercises were designed for group administration, their nature limits the size of the group to whom the exercises may be administered. With their emphasis on individual instruction during the training aspect, the MTE exercises place a heavier demand on the number of administrative personnel required and on the required level of training of the administrative personnel. Moreover, the facility and equipment requirements are more complex for MTE exercises than for the usual paper-and-pencil test. The question is whether or not such cost and convenience considerations should outweigh the benefits of an approach that is more psychometrically acceptable and is content-relevant.

Prior work (Siegel & Bergman, 1972, 1973; Siegel & Wiesen, 1977) has indicated that Navy recruits accepted the MTE approach as "fair" and favored it over a paper-and-pencil testing approach. As a personnel selection/classification test instrument, the MTEB possesses at least four important attributes: (1) it is drawn from the jobs themselves, as described in Navy documents, (2) its construction is based on accepted methods and all aspects of its administration and scoring are fully standardized, (3) the content of the

exercises is directly related to the job content, and (4) when used to test low-aptitude seamen and firemen, it selects those persons who will perform better on the job. Accordingly, a practical decision on the utility of the MTE approach to the Navy must counterbalance the cost considerations against the positive psychometric findings, the social implications, the legal implications, and the importance of proper selection/classification in the Navy.

Future Trends

Since MTEB's ability to predict Navy school grades, which represent a more proximal criterion and one that is employed quite frequently in the Navy, remains unknown, it appears that studies of MTEB's ability to predict school achievement are warranted. Within such studies, the sample should include both higher aptitude and lower aptitude personnel. Prediction of school achievement would be desirable because success in "A" school is often a prerequisite for becoming rated in the Navy.

In view of the criterion problems, which evidently exerted a negative effect on the current results, MTEB studies that employ an actual work sample criterion also seem warranted. Such an approach, while costly and time-consuming, might provide a more realistic criterion of the "true" ability of enlisted personnel after they have been in the fleet for a period of time and, accordingly, a more acceptable criterion.

It also seems that many MTE test situations could be simulated on graphic computer terminals. The training aspect might similarly be presented on a terminal through computer assisted instructional techniques. Such an approach would involve a greater initial investment, but the costs of investment might be counterbalanced by decreased administrative costs. The result might be a lower total cost of ownership. Feasibility studies seem indicated in regard to computer implementation of the MTE concept.

Finally, in view of the promise of the MTE approach, there should probably be studies of administrative and management techniques for integrating such methods into the selection/classification process for low aptitude personnel.

CONCLUSIONS

1. The MTEB and, by implication, the MTE approach possess merit for consideration within the selection/classification process for low aptitude personnel in the Navy.

2. When used in conjunction with the ASVAB to evaluate low aptitude personnel, the MTEB added to the ASVAB's ability to predict the performance of personnel assigned to fleet duties.

3. The MTE exercises seemed to have acceptable internal psychometric properties.

4. The utility of the MTE concept for Navy applications remains supportable.

RECOMMENDATION

Although the MTEB provided increases in validity coefficients over ASVAB variables, these increases were not sufficient to warrant evaluation of the cost effectiveness of using the MTEB for classification and assignment of GENDETs. Therefore, additional evaluation of the validity of the MTEB for this sample will be made using job history

variables as criteria. Cory (in press) has found job history variables to be better criteria of job performance than supervisors' marks.

Subsequent to the validation of the MTEB against job history criteria, the MTEB's usefulness and cost effectiveness for classifying and assigning GENDETs will be evaluated. As a part of this evaluation, recommendations will be made concerning future research with and the use of MTEB for GENDET personnel.

REFERENCES

Bray, D. W., & Moses, J. L. Personnel selection. Annual Review of Psychology, 1972, 23, 545-576.

Cory, C. H. Assignment of general detail personnel in the Navy: Fleet follow-up of personnel appraised in a technical classification assessment center pilot study (NPRDC Tech. Note 82-23). San Diego: Navy Personnel Research and Development Center, July 1983.

Cory, C. H. Evaluation of predictors and criteria for job performance of general detail (GENDET) personnel (NPRDC Tech. Rep.). San Diego: Navy Personnel Research and Development Center, in press.

Guilford, J. P., & Fruchter, B. Fundamental statistics in psychology and education. New York: McGraw-Hill, 1978.

O'Leary, L. R. Fair employment, sound psychometric practice, and reality: A dilemma and a partial solution. American Psychologist, 1973, 28, 147-150.

Robertson, I., & Downs, S. Learning and prediction of performance: Development of trainability testing in the United Kingdom. Journal of Applied Psychology, 1979, 64, 42-50.

Siegel, A. I., & Bergman, B. A. Nonverbal and culture fair performance prediction procedures I. Background, test development, and initial results. Wayne, PA: Applied Psychological Services, 1972.

Siegel, A. I., & Bergman, B. A. Nonverbal and culture fair performance prediction procedures II. Initial validation. Wayne, PA: Applied Psychological Services, 1973.

Siegel, A. I., & Leahy, W. R. Nonverbal and culture fair performance prediction procedures III. Cross validation. Wayne, PA: Applied Psychological Services, 1974.

Siegel, A. I., & Bergman, B. A. A job learning approach to performance prediction. Personnel Psychology, 1975, 28, 325-339.

Siegel, A. I., Pfeiffer, M. P., & Braunstein, C. Perceptions of shipboard duty--their structure and modification (NPRDC Tech. Note 77-15). San Diego: Navy Personnel Research and Development Center, August 1977.

Siegel, A. I., & Wiesen, J. P. Experimental procedures for the classification of naval personnel (NPRDC Tech. Rep. 77-3). San Diego: Navy Personnel Research and Development Center, January 1977. (AD-A035 744)

APPENDIX A

TECHNICAL PERFORMANCE CRITERION FORM FOR SEAMEN*

*The criterion forms for firemen and airmen paralleled the structure of the seaman's form

S-2

PERFORMANCE INVENTORY

Your help is needed to rate the on-job performance of the person whose name is shown in the upper left corner of this page. He is one of a group of men who have participated in research designed to evaluate a new battery of selection tests, so it is important that you complete and return this questionnaire. The information will be used for RESEARCH PURPOSES ONLY and will not affect any individual in the study.

To complete this form use the seven point rating scale that appears above the "Response" Column. For each item, circle the number which represents your choice from the scale of "1" to "7." Circle the number "1" if you think that the person performs the duty very well. Circle the number "2" if you think the person performs the duty well, and so on. If he has not performed the duty, or if he has performed the duty but you have not had the opportunity to observe his performance, rate him on how well you think he would perform the duty if he had the chance and circle the NO (Not Observed) next to the duty on the form. Make certain that only one number is circled alongside each statement. If you change your answer, erase the first circle completely and circle the number of your choice. Be sure to answer each item. For any duty which the man has not performed or a duty which you have not had the opportunity to observe his performance, rate the man on how well you think he would perform the duty and circle the NO (Not Observed) next to the duty on the form.

An example of how to complete the form is shown below.

Example

1. Very well
2. Well
3. Fairly well
4. Neither well or poorly
5. Fairly poor
6. Poorly
7. Very poorly

NO Not Observed

Duties How well does he perform or do you think he would perform the following duties:	Response							
a. Standing helm watch	1	(2)	3	4	5	6	7	NO
b. Sending and receiving code	1	2	3	4	5	6	(7)	(NO)

The person completing the form thought that the man being rated performed "standing helm watch" well. Accordingly, he drew a circle around the number "2." On the other hand, he thought that the man being rated was very poor at sending and receiving code. He circled the number "7." Moreover, he has never seen the man send or receive code. Therefore, he also circled NO (Not Observed) for "sending and receiving code."

1. Very well
2. Well
3. Fairly well
4. Neither well or poorly
5. Fairly poorly
6. Poorly
7. Very poorly

NO Not Observed

Duties How well does he perform the following duties?	Response							
1. Perform simple work related calculations	1	2	3	4	5	6	7	NO
2. Steer by compass	1	2	3	4	5	6	7	NO
3. Demonstrate knowledge of names and functions of deck and lifesaving equipment, such as anchor windlass, boat booms, accommodation ladders, winches, cranes, capstans, and davits	1	2	3	4	5	6	7	NO
4. Operate deck equipment	1	2	3	4	5	6	7	NO
5. Follow procedures for mooring and anchoring	1	2	3	4	5	6	7	NO
6. Stand variety of watches	1	2	3	4	5	6	7	NO
7. Observe safety precautions in working around tools and equipment	1	2	3	4	5	6	7	NO
8. Determine the size of a line or wire	1	2	3	4	5	6	7	NO
9. Inspect deck equipment	1	2	3	4	5	6	7	NO
10. Determine causes of failure of deck equipment	1	2	3	4	5	6	7	NO
11. Make basic log entries	1	2	3	4	5	6	7	NO
12. Maintain continuous records	1	2	3	4	5	6	7	NO
13. Read and record or use indicators of various instruments and displays	1	2	3	4	5	6	7	NO
14. Identify and describe functions of ground tackle equipment, bitts, chocks, towing pods, cleats and leadman's chains, binnacle, rudder angle indicator, engine speed indicator, and engine order telegraph	1	2	3	4	5	6	7	NO
15. Identify compasses such as standard compass, steering compass, magnesyn compass, gyro repeater and pelorus	1	2	3	4	5	6	7	NO

1. Very well
2. Well
3. Fairly well
4. Neither well or poorly
5. Fairly poorly
6. Poorly
7. Very poorly
NO Not Observed

16. Identify Navy ships and small craft	1	2	3	4	5	6	7	NO
17. Plan his own work	1	2	3	4	5	6	7	NO
18. Read diagrams and schematics	1	2	3	4	5	6	7	NO
19. Know basic rules of the road for determining burden and privilege for avoiding collision in meeting end-on, crossing, and overtaking situations, requirements of rules of the road for, and characteristics of, running and anchor lights	1	2	3	4	5	6	7	NO
20. Repair and maintain canvas and lines; make splices	1	2	3	4	5	6	7	NO
21. Serve as member of a small boat crew	1	2	3	4	5	6	7	NO
22. Identify names and characteristics of various types of ammunition	1	2	3	4	5	6	7	NO
23. Perform as sight setter, pointer, and trainer	1	2	3	4	5	6	7	NO
24. Handle and stow ammunition	1	2	3	4	5	6	7	NO
25. Tie basic knots, for example; bowline on a bight, rolling hitch, clove hitch, etc.	1	2	3	4	5	6	7	NO
26. Identify difference between bearing by degrees and bearing by points	1	2	3	4	5	6	7	NO
27. Locate and turn on ship deck and interior standing lights	1	2	3	4	5	6	7	NO
28. Compute Greenwich civil time	1	2	3	4	5	6	7	NO
29. Identify visual and sound signals for distress	1	2	3	4	5	6	7	NO
30. Identify various types of buoys	1	2	3	4	5	6	7	NO
31. Work neatly and carefully	1	2	3	4	5	6	7	NO

1. Very well
2. Well
3. Fairly well
4. Neither well or poorly
5. Fairly poorly
6. Poorly
7. Very poorly
NO Not Observed

32. Assist other technicians in maintaining deck equipment	1	2	3	4	5	6	7	NO
33. Assume adequate levels of responsibility	1	2	3	4	5	6	7	NO

34. If you were completing a NAVPERS 792 form on this man at this time, what marks should he receive? (Mark an X through the appropriate number.)

PROFESSIONAL PERFORMANCE

Not Obsd	4.0	3.8	3.6	3.4	3.2	2.8	2.6	2.0	1.0

MILITARY BEHAVIOR

Not Obsd	4.0	3.8	3.6	3.4	3.2	2.8	2.6	2.0	1.0

MILITARY APPEARANCE

Not Obsd	4.0	3.8	3.6	3.4	3.2	2.8	2.6	2.0	1.0

ADAPTABILITY

Not Obsd	4.0	3.8	3.6	3.4	3.2	2.8	2.6	2.0	1.0

APPENDIX B

FORM FOR COLLECTING CRITERION DATA ABOUT RECRUIT ADJUSTMENT

Adjustment to Navy Job Requirements

Next, please tell us about the adjustment which this man has made to the Navy and to his duty assignments.

For this part use the seven-point rating scale shown above the "Response" column. For each item, circle the number which represents your choice from the scale.

An example of how to complete the form is shown below.

Example

1. Describes man very well
2. Describes man well
3. Describes man fairly well
4. Describes man neither well nor poorly
5. Describes man fairly poorly
6. Describes man poorly
7. Describes man very poorly

The Man	Response
1. Has a large number of shipboard friends	1 **(2)** 3 4 5 6 7
2. Looks forward to returning to civilian life	1 2 3 4 5 6 **(7)**

The person completing the form thought that "the man has a large number of shipboard friends," describes the man being rated well. Accordingly, he drew a circle around the number "2." On the other hand, he thought that the statement "the man looks forward to returning to civilian life," describes the man being rated very poorly. He circled the number "7."

1. Describes man very well
2. Describes man well
3. Describes man fairly well
4. Describes man neither well nor poorly
5. Describes man fairly poorly
6. Describes man poorly
7. Describes man very poorly

The Man	Response
43. Gets along well with shipmates	1 2 3 4 5 6 7
44. Is satisfied with his job assignments	1 2 3 4 5 6 7
45. Gets assignments done on time	1 2 3 4 5 6 7
46. Does not take much time off because of sickness	1 2 3 4 5 6 7
47. Takes direction easily	1 2 3 4 5 6 7
48. Approaches his assignments conscientiously	1 2 3 4 5 6 7
49. Appears to be highly motivated on-the-job	1 2 3 4 5 6 7
50. Does not complain too much	1 2 3 4 5 6 7
51. Seems to be satisfied with his life in the Navy	1 2 3 4 5 6 7
52. Has no trouble taking orders from his supervisors	1 2 3 4 5 6 7
53. Seldom complains about working long hours	1 2 3 4 5 6 7
54. Is motivated to assume additional responsibilities	1 2 3 4 5 6 7
55. Seldom breaks rules and regulations	1 2 3 4 5 6 7
56. Seems to enjoy his work	1 2 3 4 5 6 7
57. Cooperates with others	1 2 3 4 5 6 7
58. Enjoys the opportunity to travel while in the service	1 2 3 4 5 6 7
59. Is satisfied with his work schedule	1 2 3 4 5 6 7
60. Is in good spirits most of the time	1 2 3 4 5 6 7
61. Is not too disturbed by long separations from home	1 2 3 4 5 6 7
62. Seems to be satisfied with his living conditions in the Navy	1 2 3 4 5 6 7

1. Describe man very well
2. Describes man well
3. Describes man fairly well
4. Describes man neither well nor poorly
5. Describes man fairly poorly
6. Describes man poorly
7. Describes man very poorly

	Response						
63. Seems to be satisfied with his pay	1	2	3	4	5	6	7
64. Accepts watch duty without too much complaining.	1	2	3	4	5	6	7
65. Is satisfied with the speed of advancement in the Navy	1	2	3	4	5	6	7
66. Finds duty assignments acceptable	1	2	3	4	5	6	7
67. Indicates a desire to get ahead in the Navy	1	2	3	4	5	6	7

Background Information

68. Your Name _____ Rate/Rank _____

69. How many months have you supervised this man? _____

70. To what duty station(s) has he been assigned?

_____ _____

_____ _____

71. Is he a designated striker? _____ For what rating? _____

72. Non-designated striker? _____ For what rating? _____

73. Has he taken any advancement examinations? _____

74. If so, which ones, and did he pass or fail? _____

_____

_____

75. If not, do you think that he would pass the advancement test most related to his work? _____

Page 7

DISTRIBUTION LIST

Director of Manpower Analysis (ODASN(M))
Chief of Naval Operations (OP-01), (OP-11), (OP-12) (2), (OP-13), (OP-14), (OP-15), (OP-115) (2), (OP-140F2), (OP-987H)
Chief of Naval Material (NMAT 0722), (NMAT 08)
Chief of Naval Research (Code 200), (Code 440) (3), (Code 442), (Code 448)
Chief of Information (OI-213)
Chief, Bureau of Medicine and Surgery (MED-25)
Commander of the Marine Corps (MPI-20)
Commander in Chief U.S. Atlantic Fleet
Commander in Chief U.S. Pacific Fleet
Commander Fleet Training Group, Pearl Harbor
Commander Naval Air Force, U.S. Atlantic Fleet
Commander Naval Air Force, U.S. Pacific Fleet
Commander Naval Surface Force, U.S. Atlantic Fleet
Commander Naval Surface Force, U.S. Pacific Fleet
Commander Naval Military Personnel Command (NMPC-013C)
Commander Submarine Force, U.S. Atlantic Fleet
Commander Submarine Force, U.S. Pacific Fleet
Commander Training Command, U.S. Atlantic Fleet
Commander Training Command, U.S. Pacific Fleet
Commanding Officer, Fleet Training Center, San Diego
Commanding Officer, Naval Regional Medical Center, Portsmouth (ATTN: Medical Library)
Commanding Officer, Office of Naval Research Branch Office, Chicago (Coordinator for Psychological Sciences)
Commanding Officer, Recruit Training Command (Academic Training Division)
Director, Naval Civilian Personnel Command
Director, Naval Education and Training Program Development Center Detachment, Great Lakes
Director, Training Analysis and Evaluation Group (TAEG)
Superintendent, Naval Postgraduate School
Secretary Treasurer, U.S. Naval Institute
Cammander U.S. Army Soldier Support Center, Fort Benjamin Harrison (Human Deminsion Division
Commander, Army Research Institute for the Behavioral and Social Sciences, Alexandria (PERI-ASL)
Chief, Army Research Institute Field Unit, Fort Harrison
Commander, Air Force Human Resources Laboratory, Brooks Air Force Base (Scientific and Technical Information Office)
Commander, Air Force Human Resources Laboratory, Lowry Air Force Base (Technical Training Branch)
Commander, Air Force Human Resources Laboratory, Williams Air Force Base (AFHRL/OT)
Commanding Officer Coast Guard Headquarters
Superintendent, U.S. Coast Guard Academy
Defense Technical Information Center (DDA) (12)